



# A Statistical Inference Comparison for Measurement Estimation using Stochastic Simulation Techniques

José Ismael de La Rosa Vargas, Miramontes Gerardo, Mcbride Lyle, Gilles Fleury, Marie-Eve Davoust

## ► To cite this version:

José Ismael de La Rosa Vargas, Miramontes Gerardo, Mcbride Lyle, Gilles Fleury, Marie-Eve Davoust. A Statistical Inference Comparison for Measurement Estimation using Stochastic Simulation Techniques. IEEE Transactions on Instrumentation and Measurement, 2008, Vol. 57 ((10)), pp. 2169-2180. 10.1109/TIM.2008.922098 . hal-00287379

**HAL Id: hal-00287379**

**<https://hal-centralesupelec.archives-ouvertes.fr/hal-00287379>**

Submitted on 11 Jun 2008

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A Statistical Inference Comparison for Measurement Estimation using Stochastic Simulation Techniques

José Ismael de la Rosa, *Member, IEEE*, Gerardo Miramontes, *Member, IEEE*,

Lyle E. McBride, *Senior Member, IEEE*, José J. Villa,

Signal Processing Laboratory of the Electrical Engineering Faculty

Universidad Autónoma de Zacatecas, Av. López Velarde,

98600 Zacatecas, Zac., MEXICO

Gilles Fleury, and Marie-Eve Davoust

École Supérieure d'Électricité, Department of Signal Processing and Electronics,

Plateau de Moulon, 3 rue Joliot Curie,

91192 Gif sur Yvette Cedex, FRANCE

## Abstract

The purpose of this paper is to present the comparison of different techniques for making statistical inference about a measurement system model. This comparison involves results when two main assumptions are made. First, the unknowable behavior of the probability density function (pdf)  $\wp(\mathbf{e})$  of errors, since the real measurement systems are always exposed to continuous perturbations of an unknown nature; second, the assumption that after some experimentation one can obtain sufficient information which can be incorporated into the modelling as prior information. The first assumption lead us to propose the use of two approaches which permit building hybrid algorithms; such approaches are the non-parametric bootstrap and the kernel methods. The second assumption makes possible the exploration of a Bayesian framework solution and the Monte Carlo Markov Chain (MCMC) auxiliary use to access the *a posteriori* pdf of the estimated measurand. For both assumptions over  $\wp(\mathbf{e})$  and the model, different classical criteria can be used; one uses also an extension of a recent criterion of entropy minimization. The entropy criterion is constructed on the basis of a symmetrized kernel estimate  $\hat{\wp}_{n,h}(\mathbf{e})$  of  $\wp(\mathbf{e})$ . Finally, a comparison between results obtained with the different proposed schemes is presented.

**Keywords :** Bootstrap, indirect measurement, MCMC, nonlinear regression, non-parametric pdf estimation, robust estimation.

## I. INTRODUCTION

The uncertainty characterization of a model is a whole complex problem which depends, among other characteristics, on the degrees of freedom of such a model, its behavior and its structure. The case of nonlinear

Corresponding autor, e-mail:ismaelrv@ieee.org

modelling is a special case treated at the Department of Signal Processing and Electronics (before named Department of Measurement) at the École Supérieure d'Électricité (Supélec) to deal with some problems in an indirect measurement framework (e.g. instrumentation) [18]. Tackling inverse problems according to a statistical point of view, permits the proposed methods to take advantage of the diversity and different characteristics of the collected data and all information about the collection procedure. According to previous conditions, one will have more or less information about the whole measurement system and then different considerations can be taken according to the model structure:

*Minimum of information:* In cases where limited information is available, the most natural proposition is to take a small number of hypotheses and take advantage of all information contained in the data itself using methods like the bootstrap (parametric or non-parametric) [15]–[17], [19]–[22], [39], [40], the kernel methods or non-parametric estimation [6]–[9], and all related methods [1], [11]–[14], [27], [28], [37] (see also some research work of Parzen-Rosenblatt dating of 1968).

*Well-known information:* In other cases, one may dispose of sufficient, and well-known information. If the complexity of the modelling problem can be bounded, the modelling problem can be summarized using classical parametric methods, which could lead toward analytical models [2]–[4], even if in practical problems the models are more complex than they seem.

*Maximum of information:* This last case describes ideal conditions which any researcher would like to reach, even if these conditions can be attained only after some experience with the modelling considerations and also with the treated data. In recent years, this way to tackling problems points toward the use of Bayesian methods or even ideas about data fusion. When modelling deals with complex problems but there is a maximum of information, and the problem is studied in a Bayesian framework, then the Monte Carlo procedures known as Monte Carlo Markov Chains (MCMC) [33], provide a set of tools to obtain practical solutions of the proposed models [9], [10].

The purpose when using Monte Carlo simulations like bootstrap or MCMC is to obtain the empirical probability density function (pdf) [32] which, in general, will lead to establish the different statistics about the treated data, and more specifically about the measurands of interest. The measurement modelling has been considered in various works [2]–[10]. The principal contribution of this work is presented in sections III, IV, and V where a comparison of the different schemes for estimation and uncertainty characterization is shown.

Measurement systems are formalized by two equations [4], [5] and [10]:

(i) the *Observation Equation*, which is described by the classical nonlinear regression model

$$\mathbf{y} = f(\mathbf{x}, \boldsymbol{\theta}) + \mathbf{e}, \quad (1)$$

(ii) and the *Measurement Equation* (a nonlinear function of the parametric model)

$$m_\ell = g_\ell(\boldsymbol{\theta}), \quad \ell = 1, \dots, r, \quad \text{with} \quad \mathbf{m} = \{m_\ell\}_{\ell=1}^r. \quad (2)$$

where  $\mathbf{y}$  represents the response (observed data, acquired data, or explained variables),  $\mathbf{x}$  is the vector of explicative variables or experimental protocol (for example, the measurement instants in an instrument, the sensor position, the frequencies used in the Eddy currents [5], etc.) of dimension  $n$  which is associated to data  $y_i$  ( $y_i, x_i^\top$ ),  $\boldsymbol{\theta} = \{\theta_j\}_{j=1}^p$  is the vector of functional parameters of dimension  $p$  which will be estimated by an identification procedure. The errors model the system as a random process and they are given by  $\mathbf{e}$ . The set  $\Theta_L \subset \mathbb{R}^p$  represents the admissible subset of parameters  $\boldsymbol{\theta}$ , and  $\mathbb{X}_L \subset \mathbb{R}^n$  represents the admissible domain for the explicative variables. The measurand to be estimated depends on the nonlinear mapping  $m_\ell = \mathcal{G}_\ell(f)$ , and given the pdf of parameters, one may assume that the pdf of the estimated measurands will be obtained directly by the nonlinear mapping of the pdf of parameters.

The remainder of the paper is organized as follows, the section II gives an example of a specific application of equations (1) and (2) for two measurands estimation, in this case the observed data  $\mathbf{y}$  was obtained by a simulation of the direct problem using a code of finite elements considering different grooves. The comparison of different uncertainty characterization methods begins in section III where some proposed solutions by other authors are briefly described, the comparison continues in section IV where our proposed approaches are described. The comparison finishes in section V, where the mean square error (MSE) in the sense of parameters and measurands estimation is analyzed together with the verification of some assumptions made about the noise nature by inspecting the residuals. Finally, some concluding remarks are given in section VI

## II. ESTIMATION OF GROOVE DIMENSIONS BY RFEC

The final objective in this paper is to present a practical evaluation perspective when using some different approaches to make statistical inferences in a general context of measurement modelling or measurand estimation. The Remote Field Eddy Current (RFEC) inspection technique<sup>1</sup> is used for dimensioning grooves that may occur in ferromagnetic conductive pipes (this is a practical measurand estimation problem, however, in this study only simulated data is used). This task involves the depth  $d$  and the length  $l$  estimation of corrosion grooves from measurements of a picked-up coil signal phase  $y_i$  at different positions  $x_i$  closest to the defect (measurand estimation problem presented in [4], [5] and [10]). The dimensions of the groove to be estimated are linked to the parametric model (eq. (1)) through a polynomial function (eq. (2)). Previous knowledge of the relationship between the groove parameters (length and depth) and the observed data (detector phase) is needed.

<sup>1</sup>Eddy Currents were discovered by the French physicist Jean Bernard Léon; these currents are induced in metallic masses under the effect of variable magnetic fields.

In order to make statistical inferences about parameters and groove estimators, the order of results presentation, according to the methodologies used, is the following:

- (i) statistical characterization using mainly bootstrap methods,
- (ii) statistical characterization using MCMC methods, and
- (iii) statistical characterization using kernel methods together with bootstrap methods and Monte Carlo.

The use of bootstrap techniques to obtain the statistical information about parameters and measurand estimates is made with the aim of quantifying the performance of some different proposed estimators; additionally, such techniques can be used to implement a complete scheme of model selection [36] (this is not the aim in this work, see [3], [4]). In such a case, a modelling error vector (e.g. errors between the observed data and the nonlinear model) whose distribution is unknown is used.

In the second proposition, the obtained MCMC results are compared with those obtained by using a Primitive Monte Carlo (PMC) scheme. In this last scheme, the hypothesis of error normality is taken into account ( $e_i \sim \mathcal{N}(0, \hat{\sigma}_{ML}^2)$ ), where  $\hat{\sigma}_{ML}^2$  is an unbiased variance estimate (the distribution is assumed to be known). The PMC scheme builds a likelihood empirical population by using the nonlinear least squares (NLS) estimator. In such a case, the modelling errors plus acquisition errors (e.g. errors between the observed data and the nonlinear model) are assumed to belong to a Gaussian distribution with unknown variance<sup>2</sup>.

The final proposition is based on the use of bootstrap techniques [7] and the Maximum-Entropy Estimation (MEE) strategy (based on the kernel methods) to obtain statistical information about parameters and measurement estimates. As the modelling, and acquisition errors distribution is supposed unknown, we use bootstrap to resampling the vector of errors (i.e. sampling from the empirical distribution). We compare the normality hypothesis of errors ( $e_i \sim \mathcal{N}(0, \hat{\sigma}_{ML}^2)$ ) with those obtained above by bootstrapping errors. Additionally, we use another recent Monte Carlo strategy known as Monte Carlo by Latin Hypercube Sampling (MCLHS) [23].

Figure 1:

Figures 1 and 2 illustrate the physical problem of groove estimation and the device which is used with the aim to collect the observed data  $\mathbf{y}$ . The sensor is introduced into the pipe and moved along the sense of  $x$  axe with the range of 156 mm of longitudinal distance, the phase difference  $y_i$  between voltage of the excitation coil and the receiver coil is measured according to position  $x_i$  of the sensor (at uniform distances of 1.33 mm). The distance between both coils ( $2L$ ) is selected such that it satisfies the condition of far-field ( $2L >$  which is two times the diameter of the pipe).

Figure 2:

<sup>2</sup>The variance is assumed to be constant and must be estimated using the Maximum Likelihood (ML) estimator.

The analytical function used to model the observed data  $\mathbf{y}$  [3], [4] is (3), and it takes into account the asymptotic behavior of data and its symmetry (see Figure 3),

$$\begin{aligned} f(\mathbf{x}, \boldsymbol{\theta}) = & \theta_1(\arctan(\theta_2(\mathbf{x} + \theta_3 + L)) - \arctan(\theta_2(\mathbf{x} - \theta_3 + L)) \\ & + \arctan(\theta_2(\mathbf{x} + \theta_3 - L)) - \arctan(\theta_2(\mathbf{x} - \theta_3 - L))) \end{aligned} \quad (3)$$

where  $p = 3$ ,  $L$  represents the half distance between both coils (constant  $L = 17.5$  mm). The length of vectors  $\mathbf{x}$  and  $\mathbf{y}$  is  $n = 118$ . The measurement goal is to estimate the size of the defect (i.e. groove) from the knowledge of detector phase. The measurement quantity is expressed as a function of the optimal parameters of the model as shown also in [4].

Once the parameters estimate has been obtained, the final process is to estimate the dimensions  $d$  and  $l$  of the defect. Thus, the bilinear polynomial function  $g$  previously selected and optimally adapted for the requested problem [5] is,

$$\begin{aligned} \hat{d} = g_d(\boldsymbol{\theta}) &= \sum_{k,h} c_{kh}^d \theta_k \theta_h + \sum_k c_k^d \theta_k \\ \hat{l} = g_l(\boldsymbol{\theta}) &= \sum_{k,h} c_{kh}^l \theta_k \theta_h + \sum_k c_k^l \theta_k \end{aligned} \quad (4)$$

These polynomial relations were proposed during the simulation of some different grooves (i.e. different defect dimensions). The coefficients  $(c_{kh}^d, c_k^d)$  and  $(c_{kh}^l, c_k^l)$  were obtained using the optimal technique of Least Squares. These procedures can also be used for a certain number of related measurand estimates. According to the candidate model, an analytical form for the measurand estimation vector is  $\hat{\mathbf{m}} = [\hat{d}, \hat{l}]^\top$  viewed as a function of  $\boldsymbol{\theta}$ .

The first objective in the comparison is to quantify the degree of accuracy of each method taking into account the final modelling goal. Such a comparison also touches other related methods which give more exhaustive information. A distance measure to make comparisons between the different results given by the compared methods is the mean square error (MSE) function:

(a)  $\text{MSE}_{\boldsymbol{\theta}}$  obtained according to the fitting function sense,

$$\text{MSE}_{\boldsymbol{\theta}} = \frac{1}{n} \sum_{i=1}^n \left( y_i - f(x_i, \hat{\boldsymbol{\theta}}) \right)^2, \quad (5)$$

(b)  $\text{MSE}_{\mathbf{m}}$  obtained using the estimated measurands and the measurands of reference (known):

$$\text{MSE}_{\mathbf{m}} = \frac{1}{2} \sum_{j=1}^2 (m_{rj} - \hat{m}_j)^2. \quad (6)$$

All calculations were dedicated only to one reference groove (pair of measurements), where the depth is  $d = 1.66$  mm and the length is  $l = 4$  mm; then the measurand vector is assigned as  $\mathbf{m}_r = [1.66, 4.0]^\top$ . Additionally, the results obtained can be compared in the same way for different groove dimensions.

### III. ALREADY PROPOSED SOLUTIONS

In most cases of multidimensional measurement estimation, the methods of Monte Carlo are effective tools which can be used, however, in such cases some hypotheses about the errors must be made, and sometimes

such hypotheses are subject to be verified (as in the case of the method PMC). In order to make statistical inferences about the groove estimators when using RFEC, Brahim [4] has already proposed to use the Monte Carlo Latin Hypercube Sampling (MCLHS) [24] and [25]. The MCLHS is a procedure which uses a variance reduction technique (VRT). For such technique, the number of simulations  $N$  could be significantly reduced. Keramat [23] has developed some ideas about VRT techniques which are used in the statistical analysis and performance optimization framework of electronic circuits manufacture. He has studied the asymptotic behavior of MCLHS methods and a modified version of this method has also been proposed to perform over the time of computation, this is the MCMLHS method. However, in the Keramat's work it is not well specified the effective way to obtain  $N$  such that the statistical characterization would be precise with respect to that obtained with the MCP method.

On the other hand, it has been specified that if  $N$  is sufficiently small, the MCLHS and MCMLHS methods will not meet the condition of efficiency, this means that variance is weak, but the mean value could be very accurate. Then, if the mean estimate is only the parameter of interest, the time of simulation will be reduced. However, in the cases where the whole pdf characterization is needed the MCLHS and MCMLHS could be used, if the proposition 3.1 by Keramat [23] is respected. In this last case, it is important to establish some relationship between  $N$  and the efficiency in function of a lower bound which is also called "the worst case bound":

$$\gamma_{LB} = \begin{cases} 1 & \text{si } N = n = 1, \\ \frac{N-1}{N} & \text{if } n > 1 \text{ and if } N > 1, \end{cases} \quad (7)$$

and in a more general way, this lower bound satisfies the following equation:

$$\gamma_{LB} = \frac{N}{N + \left(\frac{N}{N-1}\right)^n \frac{n(n-1)}{2}}, \quad \text{if } n > 1. \quad (8)$$

In this case, the efficiency condition  $\gamma_{LB} \rightarrow 1$  for a largest  $N$ , and then the theorem 3.5 proposed by Keramat [23]:  $\text{Var}_{\text{PMC}} \geq \text{Var}_{\text{LHS}}$ , is well respected. Then, at this stage the next question remains without answer: which is the adequate value for  $N$  in order to obtain sufficient (asymptotically) accuracy about the pdf which will be characterized?

Obviously, in the present work such a question is given only for reflection and indeed future efforts could be made trying to respond. In terms of comparison the value  $N = 5000$  was heuristically chosen to achieve a report with respect to the bootstrap in which case the number of proposed simulations is  $B = 5000$  (will be detailed in next section), and also with respect to PMC method with a superior number of simulations,  $N = 50,000$ . The hypotheses assumed for both methods MCLHS and PMC deal with Gaussian behavior, where the variance is assumed constant and unknown, thus the unbiased Maximum Likelihood estimator is

used to obtain  $\sigma^2$ :

$$\hat{\sigma}_{ML}^2 = \frac{1}{n-p} \sum_{i=1}^n (y_i - f(x_i, \boldsymbol{\theta}))^2. \quad (9)$$

Under Gaussian hypotheses the NLS<sup>1</sup> estimator ( $e_i \sim \mathcal{N}(0, \hat{\sigma}_{ML}^2)$ , where  $\hat{\sigma}_{ML} = 0.0139$ , see above equation) is a good choice and thus it has been used. The results of simulation were obtained using a SUN Sparc 5 station (330 MHz, 256 Mbytes of RAM), and the simulation platform was MATLAB version 5.2. A modification in the number of iterations into the Levenberg-Marquardt algorithm lead to the NLS<sup>2</sup> estimator, the same computational equipment has been used and the platform of simulation was changed by MATLAB version 6. Table I illustrates the comparison between some obtained statistics using PMC vs. MCLHS schemas, when the number of iterations is increased the schema PMC-NLS<sup>2</sup> lead to a most concentrated empirical density (pdf), and the statistics about measurements are the best approximated with respect to the reference (however, the time of simulation could be prohibitive).

Table 1:

Brahim [4] propose to use a second approach which has been developed in a deterministic framework [26]. In this case, some procedures based on the bounded errors are analyzed together with the set of observed data  $\mathbf{y}$ . The uncertainty characterization about the measurement system is no more made according to a probabilistic point of view, such uncertainty is made according to some ensemble or set manipulations and an analysis of complex intervals. These two tasks are driven by using a type of algorithms called SIVIA (Set Inversion Via Interval Analysis) and ImageSP (Image Sub-Paving evaluation). This last approach has been used taking advantage of collaboration with the Signals and Systems Laboratory (LSS) at Supélec [4].

#### IV. NEW PROPOSED SOLUTIONS

##### A. The mainly bootstrap approach

This section deals with the use of bootstrap methods for statistical characterization of parameters according to the model (3) and for measurands  $\widehat{\mathbf{m}}$  ( $\widehat{d}$  and  $\widehat{l}$ ) by using the relation (4). The hypotheses about the errors, in this case, are limited to some few assumptions taking into account that the modelling and acquisition errors (i.e. errors between the observed data  $\mathbf{y}$  and the function model  $f(\cdot)$ ) have an unknown distribution  $\wp(\mathbf{e})$ . All the proposed bootstrap procedures, were simulated using the value  $B = 5000$  times,  $B$  defines the total number of bootstrap simulations. Table II shows some statistics obtained for  $\boldsymbol{\theta}$  and for  $\mathbf{m}$  when using the methods described in [8], these are the weighted bootstrap compared versus the NLS estimator. Similarly, Table III illustrates also some statistics obtained using the weighted bootstrap method, but in this case the robust or biased bootstrap estimators (Biased Bootstrap with Quadratic criterion - BBQ) and (Biased Bootstrap with quasi-Huber criterion - BBH) are used.

Table 2:



In the use of robust estimators (BBQ and BBH), it is important to permit a good adjustment of data which could be contaminated by a non-Gaussian noise or outliers. Figure 5 in [8] illustrates some fittings of data using the model (3), in such a case the mean values of parameters were used after the simulation procedure of the BBH estimator where the weighted bootstrap resampling was also used. In the same figure the circles represent the observed data under analysis (such data were generated by using a finite element code for a groove with  $d = 1.66$  mm and  $l = 4$  mm). On the other hand, Tables II and III show some statistics when using, in both estimators, the weighted bootstrap techniques with non-centered and centered residuals, bootstrap-NLS<sup>1</sup> (i.e. the mean value  $\bar{r}$  is eliminated from equation (11) in [8]). The statistical behaviors obtained when using the NLS and BBQ estimators are very similar. In the case of the weighted bootstrap with centered residuals, bootstrap-NLS<sup>2</sup>, the obtained statistics from  $d$  and  $l$  for both estimators NLS and BBQ are also similar but in this case more precise (see Tables II and III), the visualized dispersion is small (bias between reference values and the mean values, and variance reduction is well remarkable). The centering factor permits robustness into the optimization procedure.

The BBH estimator gives more precision about the measurement estimation, the obtained populations are less dispersed, and thus, the obtained statistics have smallest values than the statistics obtained when using the NLS and BBQ estimators (see also Tables II and III).

Table 3:

Figure 3:

It is important to remark that the analyzed data do not appear to be contaminated by outliers, and thus, it is not surprisingly the great likelihood remarked about the statistical behaviors obtained when using the BBH and NLS estimators (see Table III). On the worst case of outliers presence, the robust estimators can be used as the best alternative illustrated in [7] (in the efficiency sense). However, if the time is a constraint, then classical estimators are the best option, for example, the time of convergence needed by the bootstrap-BBH schema after 5000 simulations was of an order of 58.21 minutes, whereas the bootstrap-NLS and bootstrap-BBQ time of simulation remains in the range of 6.67 to 12.07 minutes for the same number of simulations. The precision level of  $\hat{d}$  and  $\hat{l}$  estimates depends on the function model  $f(\cdot)$ , but also of an adequate choice of the polynomial functions  $g_\ell(\cdot)$ ,  $\ell = 1, \dots, r$ . All times of simulation presented along this paper were obtained using the same computational equipment (SUN Sparc 5 station of 330 MHz, 256 Mbytes of RAM, and MATLAB 5.2 version).

### B. Characterization using MCMC

Other proposed approach is constructed on the basis of a bayesian framework, in such approach the MCMC tools are necessary to obtain the full statistical characterization. Some additional hypotheses about the errors distribution are taken into account by regarding the error sample distribution (which is non asymptotical),

for such considerations the obvious estimator is the NLS since the sample distribution seems to be centered (mean value almost zero). In fact, the classical hypotheses about Gaussian behavior of errors, with variance  $\sigma^2$  unknown but constant are reconsidered. In practice, the errors distribution of course is unknown, but some times it is asymptotically Gaussian. In this case, it is important to verify if the hypotheses made are justified or valid (see next section § V; in this sense the formulated hypotheses to deal with the PMC and MCLHS procedures will be tested).

For simplicity of assumptions, the prior probability of parameters is considered unknown, this means that the most adequate prior probability function is the noninformative or uniform (i.e. it could be also a normal prior as illustrated by Sandu [34], [35]). The probabilistic model function used for the variance hyper-parameter is thus the inverse gamma probability density function, that is  $\sigma^2 \sim \mathcal{IG}(\alpha, \beta)$ .

The approaches detailed in [10] are used, these are the Gibbs simulation procedure and the hybrid Gibbs Metropolis-Hastings procedure. If the assumed hypotheses about the errors are valid, the performance comparison could be made in a similar framework of the bootstrap proposed methods but in this case using auxiliary tools such as MCMC, and also equivalent framework with respect to PMC and MCLHS (see section § III). In this case, The Bayesian framework is thus equivalent with the maximum likelihood principle and so the least squares estimator basis.

### B.1 The Gibbs approach

In the most simple case of the Gibbs adaptation, the results depend on the prior modelling of pdf functions about the parameters and hyper-parameters of interest, thus the posterior densities are obtained under the basis of  $\varphi(\boldsymbol{\theta}|\mathbf{y}) \propto \mathcal{L}(\mathbf{y}|\boldsymbol{\theta})$ , in such a case the prior distribution of parameters has an uniform behavior over the interval  $[0, 1]$ . When the MCMC is used, it is necessary to consider a “burn in time” (e.g. some initial simulations must be discarded) of simulation, for this reason the total number of simulations made is 6000 simulations for a single Gibbs chain, improving the time of convergence by using a weighted bootstrap technique, this simulation scheme will be labeled by (MCMC<sup>1</sup>), The first 1000 simulations are wasted (the burn) and the remaining 5000 simulations were used to obtain the whole statistical populations figured out in Table IV.

On the other hand, a modification of the previous scheme is suggested. In this new proposition, the convergence improvement of the Gibbs sampling is lead by recent algorithms of perfect simulation (see references [29] and [31], where MCMC<sup>2</sup> uses the coupling concept of two chains), this schema has been simulated 4500 times (the first 500 simulations have been discarded, even if theoretically there is no problem of burn). Generally and intuitively, after comparison with PMC it seems that the MCMC<sup>1</sup> and MCMC<sup>2</sup> approximated statistics are weak on accuracy, even if this statistics could be identified into the neighborhood of the obtained statistics by PMC-NLS<sup>1</sup> approach. However, the simulation time has decreased considerably.

Table 4:

### B.2 An hybrid Gibbs M-H approach

A second class of MCMC approaches was also considered, two approaches based on an hybrid conjunction of MCMC are also used. In the sampling procedure is substituted the likelihood function  $\mathcal{L}(\mathbf{y}|\boldsymbol{\theta})$ , previously used according to the Gibbs sampling, by a new function that has been proposed by Pázman [30]. This last likelihood function takes into account the nonlinearities of parameters and also the nonlinearities mapped to the measurands  $d$  et  $l$ . The proposed matrix  $\Delta$  has the assignment  $\Delta = \lambda \mathbf{K}(\boldsymbol{\theta})$ , with an heuristic value for  $\lambda = 3.2$  and  $\mathbf{K}(\boldsymbol{\theta})$  obtained according to the Cholesky decomposition of  $\mathbf{F}(\boldsymbol{\theta})^{-1}$  which is the inverse of the Fisher information matrix [10]. The number of realized simulations of the Gibbs M-H scheme was 6000 for a single chain (G-M-H<sup>1</sup>) (the first 1000 were discarded and the remaining 5000 were used to construct the statistical populations).

In the case of the scheme which uses a double chain to improve convergence or Gibbs M-H by perfect simulation (MCMC<sup>4</sup>), as in the case of Gibbs perfect sampling 4500 simulations were calculated. Figure 3 shows some data fitting when using the mean value of parameters (trimmed line) and according to the MCMC<sup>4</sup> sampling scheme. In the same figure, it is also illustrated the comparison versus the fitting given when using the PMC-NLS<sup>2</sup> scheme, there are also some differences between the MCMC<sup>2</sup> and MCMC<sup>4</sup> schemes. It is clearly argued that Gibbs Metropolis-Hastings scheme captures better the nonlinearity of the first modelling function (3) according to the Pázman's work.

The final acceptance percent obtained using the MCMC<sup>3</sup> scheme diminishes toward 18.35%, this is driven by changing the  $\lambda$  values. Thus, the rejection percentage of samples is of the order of 81.65%. If the mean time of simulation is compared locally, it has been observed that the MCMC<sup>3</sup> and MCMC<sup>4</sup> schemes are slowest than the MCMC<sup>1</sup> and MCMC<sup>2</sup> schemes (The Gibbs sampling scheme is about 30 times faster than the PMC-NLS<sup>1</sup> scheme), however, their accuracy level is far better. Other example of time comparison is that obtained when using the PMC-NLS<sup>1</sup> and PMC-NLS<sup>2</sup> approaches, the elapsed time goes from 1.42 hours to 7.32 hours respectively. The mean simulation times for MCMC<sup>3</sup> and MCMC<sup>4</sup> schemes varies from 46 minutes to 72 minutes. Obviously, the accuracy level of the estimates  $\hat{d}$  and  $\hat{l}$  can be compared with the obtained one using PMC-NLS, moreover accuracy is also related to the adequate choice of  $f(\cdot)$  and  $g_\ell(\cdot)$ , where  $\ell = 1, \dots, r$ .

Table 5:

Figure 4:

### C. Characterization using kernel and bootstrap methods

This section illustrates another proposed approach which is based on the Minimum-Entropy Estimation (MEE) concepts and the simultaneous use of the bootstrap simulation methods [6], [7] (as seen in § IV-A). The final objective is the same, to obtain a complete statistical characterization of the MEE parameter estimators and

$d$  and  $l$  estimators. As in the case of only bootstrap, here also one has limited assumptions and to smooth the empirical distribution of errors the use of bootstrap methods is appropriate, the model and the acquisition errors are stochastic variables independent and identically distributed (s.v.i.i.d), and its distribution is also assumed to be of unknown structure. Other additional consideration is that such variables are centered in some sense. As it has been seen previously, the empirical distribution sampling of residuals can be lead mainly by the weighted bootstrap methods. Indeed, the hypotheses assumed could be more specific and so, for simplicity the classical Gaussian assumption may be considered ( $e_i \sim \mathcal{N}(0, \hat{\sigma}_{ML}^2)$ ), as in PMC-NLS schemes in section § III (PMC-NLS<sup>1</sup> and PMC-NLS<sup>2</sup>).

In fact, the Figure 4 illustrates a visual comparison between modelling data and fitting when using three different MEE estimators studied previously (three nonparametric-kernel methods) [7]. Such a comparison is made in contrast to the NLS estimator (in all cases the mean value of parameters is used). On the other hand, Table VI shows also some statistical values obtained for  $\theta$ ,  $d$  and  $l$ , according to the three MEE estimators. The statistics with the poorest accuracy are those obtained with the MEE<sup>1</sup> (Gaussian kernel) estimator, where it is perceived the largest bias. In the case of the three estimators, the statistical population of parameters and measurands is generally more concentrated than PMC-NLS, so the variance has a lower bound which is weaker than the Cramer-Rao bound.

Figure 5:

According to results obtained when simulating the MEE<sup>2</sup> estimator (with cosine weighted functions as kernel), Table VI depicts the statistics which are included in the neighborhood of three estimators. The remaining estimator is MEE<sup>3</sup> (which uses a Hilbet Kernel), this last estimator performs better among the three proposed. Comparing all statistical populations of estimators it is evident that the most accurate are those obtained when simulating the MEE<sup>3</sup> estimator.

The different proposed MEE schemes generally lead to good statistical characterizations and estimates are relatively accurate, as it has been seen in examples presented in [7]. In the estimation case of the defects by using the RFEC technique, once again, it is confirmed that results offered by the estimator MEE<sup>3</sup> give a good performance. The statistical characterization of  $d$  and  $l$  according to the scheme bootstrap-MEE<sup>3</sup> is consistent (considering the size of  $\mathbf{y}$ ), the calculated mean values suggest the existence of an agreement of the mean values and the level of the parameters and measurands population dispersions.

Finally, according to the calculation time, the MEE estimators do not offer the best performance. For example, the first parameters estimate and measurements estimate when using the three proposed MEE estimators need the following times: 69.9540 sec with MEE<sup>1</sup>, 19.7344 sec with MEE<sup>2</sup> and 21.6249 sec with MEE<sup>3</sup>. The best performance offered is given by the estimator MEE<sup>2</sup> (for simulation purposes the use of MEX files was avoided, however its inclusion could improve the calculation time performance). On the other hand, when comparing the above times vs. times offered by the NLS estimator (0.0812 sec), it is clear that

the efficiency of least squares method is very good, and its only disadvantage is its limited robustness.

Table 6:

Figure 6:

Figure 7:

## V. NORMALITY ANALYSIS AND RESULTS JUSTIFICATION FOR THE MSE DISTORTION

In a first analysis, the estimates of parameters and measurands ( $\hat{\boldsymbol{\theta}}$ ,  $\hat{d}$  and  $\hat{l}$ ) are compared. This is done with the aim of obtaining two different distortion measures according to the mean square error (MSE) sense:

1. the approximation quality between  $f(\mathbf{x}, \hat{\boldsymbol{\theta}})$  and  $\mathbf{y}$  (see equation (5)), and
2. the approximation quality between  $\mathbf{m}_r$  and  $\hat{\mathbf{m}}$  (see equation (6)).

In a second analysis stage, the distortion measures of the mean square error (MSE) are also obtained, but in this case the parameters and measurands estimated mean values are used. Tables VII and VIII show the calculations of the MSE obtained for both distortion measures, and for the estimate and the mean estimated values of parameters and measurands.

Tables 7, 8, 9:

The final analysis stage consists on the hypotheses evaluation about the Gaussian behavior of errors, the first two empirical statistical moments of errors are calculated:

$$\hat{m}_e = \frac{1}{n} \sum_{i=1}^n e_i(\hat{\boldsymbol{\theta}}), \quad (10)$$

and

$$\hat{v}_e = \frac{1}{n-p} \sum_{i=1}^n \left( e_i(\hat{\boldsymbol{\theta}}) - \hat{m}_e \right)^2, \quad (11)$$

and, if the following constraint with respect to the mean value is not satisfied:

$$|\hat{m}_e| > \text{Test}_1, \quad \text{Test}_1 = 2\sqrt{\frac{\hat{v}_e}{n}}, \quad (12)$$

then the hypothesis or assumption of zero mean will be retained.

On the other hand, the normality hypotheses about the errors can also be verified in different manners [38]. For simplicity, the property of a normal law will be retained by using notions about the skewness and kurtosis coefficients. The skewness coefficient is defined by the following expression:

$$\gamma_1 = \frac{E \left\{ (e - E \{e\})^3 \right\}}{\sigma^3}, \quad (13)$$

and for the case of symmetric distributions,  $\gamma_1 = 0$ . One approximation to such a coefficient is given by the following estimator:

$$\hat{\gamma}_1 = \frac{\frac{1}{n} \sum_{i=1}^n \left( e_i(\hat{\boldsymbol{\theta}}) - \hat{m}_e \right)^3}{\hat{v}_e^{\frac{3}{2}}}. \quad (14)$$

The kurtosis coefficient is defined by the following expression:

$$\gamma_2 = \frac{E \left\{ (\mathbf{e} - E \{ \mathbf{e} \})^4 \right\}}{\sigma^4} - 3, \quad (15)$$

and one approximation of  $\gamma_2$  is given by the following estimator:

$$\hat{\gamma}_2 = \frac{\frac{1}{n} \sum_{i=1}^n \left( e_i(\hat{\boldsymbol{\theta}}) - \hat{m}_e \right)^4}{\hat{v}_e^2} - 3. \quad (16)$$

By definition, the kurtosis coefficient  $\gamma_2 = 0$  when the distribution is Gaussian; then the equality  $\gamma_1 = \gamma_2 = 0$  is expected in general. According to the normality hypothesis,  $\hat{\gamma}_1$  and  $\hat{\gamma}_2$  have the following mean equations,

$$m_1 = 0, \quad m_2 = -\frac{6}{n+1},$$

and their respective variances are:

$$v_1 = \frac{6(n-2)}{(n+1)(n+3)}, \quad v_2 = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}.$$

Then, the normality hypothesis will be rejected if the following conditions are satisfied:

$$\text{Test}_2 = \frac{|\hat{\gamma}_1|}{2\sqrt{v_1}} > 1, \quad (17)$$

and if

$$\text{Test}_3 = \frac{|\hat{\gamma}_2 - m_2|}{2\sqrt{v_2}} > 1. \quad (18)$$

Other types of analysis can also be made, those for stationarity and independence. The respective tests are also presented by Walter and Pronzato [38].

Table 10:

The results obtained for the different tests are presented in Table X. The normality errors tests are positive for almost all estimators used (e.g. the  $\text{Test}_1 > |\hat{m}_e|$  and the  $\text{Test}_2$  and  $\text{Test}_3$  give values less than 1). In fact, this is the reason why the  $\text{MSE}_{\boldsymbol{\theta}}$  is minimum when the  $\text{NLS}^1$ ,  $\text{NLS}^2$  and  $\text{BBQ}$  estimators are used (see Tables VII and VIII). According to the  $\text{MSE}_{\boldsymbol{\theta}}$  the above estimators are optimal in the groove dimensioning problem. On the other hand, according to the  $\text{MSE}_{\mathbf{m}}$ , it is well minimized when the  $\text{MEE}$  estimators are used, and particularly the  $\text{MEE}^2$  and  $\text{MEE}^3$  performance is excellent. The results, when the empirical mean is estimated by using the  $\text{MEE}^3$  estimator, confirm its great performance (see Table VIII).

## VI. CONCLUSIONS AND FUTURE WORK

The uncertainty characterization techniques of measurement systems proposed in this work depend on the following considerations: if one disposes of limited knowledge about distribution errors and with a limited (small) size of data, it is recommended to use:

- bootstrap techniques for their simplicity and easy use,
  - kernel based on cosine weighted functions (robust MEE estimator) and bootstrap,
- if one disposes of prior knowledge, it is recommended to use interaction between conditional probabilities according to the Bayesian framework and find the posterior probability density:
- using sampling MCMC techniques, and
  - improvement of convergence using resampling techniques and the coupling from the past ideas.

Perspectives: the asymptotical analysis of MEE estimators is nowadays an open problem, the sampling MCMC improvement by using slice sampling is another technique which will be interesting to explore in future works.

#### ACKNOWLEDGEMENTS

Many thanks to PROMEP of Mexico, this work was partially supported by the Mexican Program for Professor's Technical Improvement (PROMEP) under Grant UAZAC PTC 24-103.5/03/1127, and the facilities offered by CITEDI-IPN during the stage invitation in December 2004.

#### REFERENCES

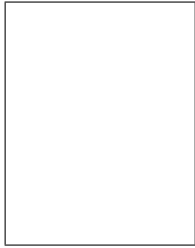
- [1] A. Berline, and L. Devroye, "A comparison of kernel density estimates," *Publications de l'Institut de Statistique de l'Université de Paris*, Vol. 38(3), pp. 3-59, 1994.
- [2] S. Brahim-Belhouari, and G. Fleury, "Probability distribution in nonlinear estimation -a measurement dedicated approach-," *Proc. 9th IEEE workshop on Statist. Signal and Array Proc.*, pp. 395-398, Portland 1998.
- [3] S. Brahim-Belhouari, M. Kieffer, G. Fleury, L. Jaulin, and E. Walter, "Model selection via worst-case criterion for nonlinear bounded-error estimation," *IEEE Trans. Instrum. Meas.*, Vol. 49, 3, pp. 653-658, June 2000.
- [4] S. Brahim-Belhouari, *Choix de structures de modèles pertinentes pour un problème d'inversion*, Disserted Ph.D. thesis, Université Paris XI-SUPELEC, Sept. 2000.
- [5] M.-E. Davoust, G. Fleury, and J. Oksman, "A parametric estimation approach for grooves dimensioning using remote field eddy current inspection," *Research in Nondestructive Evaluation*, Vol. 11, pp. 39-57, 1999.
- [6] J. I. De la Rosa, and G. Fleury, "On the Kernel selection for Minimum-Entropy estimation," *Proc. of the IEEE Instrumentation and measurement Technology Conference*, Vol. 2, pp. 1205-1210, Anchorage, AK (USA), 21-23 May 2002.
- [7] J. I. De la Rosa, G. Fleury, and M.-E. Davoust, "Minimum-Entropy, pdf approximation and Kernel selection for measurement estimation," *IEEE Trans. Instrum. Meas.*, Vol. 52, No. 4, pp. 1009-1020, August 2003.
- [8] J. I. De la Rosa, and G. Fleury, "Bootstrap methods for a measurement estimation problem," *IEEE Trans. Instrum. Meas.*, Vol. 55, No. 3, pp. 820-827, June 2006.
- [9] J. I. De la Rosa, *Estimation de la densité de probabilité d'une mesure dans un cadre non-linéaire, non-gaussien*, Disserted Ph.D. thesis, Université Paris XI-SUPELEC, Nov. 2002.
- [10] J. I. De la Rosa, G. Fleury, and S. E. Osuna, "Density estimation for measurement purposes and convergence improvement using MCMC," *IEEE Instrumentation and measurement Technology Conf. IMTC-2003*, Vol. 1, pp. 478-483, Vail, Colorado (USA), 20-22 May 2003.
- [11] L. Devroye, "A note on the usefulness of superkernels in density estimation," *The Annals of Statistics*, Vol. 20, pp. 2037-2056, 1992.

URL : [//jeff.cs.mcgill.ca/~luc/](http://jeff.cs.mcgill.ca/~luc/)

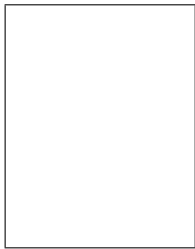
- [12] L. Devroye, "The double kernel method in density estimation," *Annales de l'Institut Henri Poincaré*, Vol. 25, pp. 533–580, 1989.
- [13] L. Devroye, and A. Krzyżak, "On the Hilbert kernel density estimate," *Statistics and Probability Letters*, Vol. 44, pp. 299–308, 1999.
- [14] L. Devroye, "Universal smoothing factor selection in density estimation: theory and practice," *Test*, Vol. 6, pp. 223–320, 1997.
- [15] N. DiCiccio, and J. P. Romano, "A review of Bootstrap confidence intervals (with discussion)," *J. Royal Statist. Soc., Series B*, Vol. 50, 3, pp. 338–370, 1988.
- [16] B. Efron, "Bootstrap methods : Another look at the Jackknife," *Annals of Statist.*, Vol. 7, pp. 1–26, 1979.
- [17] B. Efron, "Better Bootstrap confidence interval," *J. of the American Statist. Assoc.*, Vol. 82, pp. 171–185, 1987.
- [18] G. Fleury, "Optimal nonlinear modeling and reparametrization," *IEEE International workshop on Intelligent Signal Proc.*, pp. 72–76, sept. 1999.
- [19] D. A. Freedman, "Bootstrapping regression models," *The Annals of Statist.*, Vol. 9, pp. 1218–1228, 1981.
- [20] P. Hall, and B. Presnell, "Applications of intentionally biased Bootstrap methods," *Documenta Mathematica*, Extra Vol. ICM III, pp. 257–266, 1998.
- [21] P. Hall, and B. Presnell, "Biased Bootstrap methods for reducing the effects of contamination," *J. Royal Statist. Soc., Series B*, Vol. 61, 3, pp. 661–680, 1999.
- [22] D. V. Hinkley, "Bootstrap methods," *J. Royal Statist. Soc., Series B*, Vol. 50, 3, pp. 321–337, 1988.
- [23] M. Keramat, *Analyse statistique et optimisation du rendement de fabrication des circuits électroniques*, Disserted Ph.D. thesis, Université Paris XI-SUPELEC, Jan. 1998.
- [24] M. Keramat, and R. Kielbasa, "Theoretical study of Latin Hypercube Sampling in Monte Carlo estimation of integrals," *Seminars-97 Process Statistics in Medical Applications*, pp. 99–125, University Paris V, France, 1997.
- [25] M. Keramat, and R. Kielbasa, "Modified Latin Hypercube Sampling Monte Carlo (MLHSMC) estimation of average quality index for integrated circuits," *Analog Integrated Circuits and Signal Processing*, Vol. 19, No. 1, pp. 87–98, april 1999.
- [26] M. Kieffer, *Estimation ensembliste par analyse par intervalles - Application à la localisation d'un véhicule*, Disserted Ph.D. thesis, Université de Paris XI, Jan. 1999.
- [27] C. M. Loader, "Bandwidth selection: classical or plug-in?," *The Annals of Statistics*, Vol. 27, No. 3, pp. 415–438, 1999.
- [28] E. Masry, "Probability density estimation from sampled data," *IEEE Trans. on Information Theory*, Vol. IT-29, No.5, pp. 697–709, Sept. 1983.
- [29] J. Møller, "Perfect simulation of conditionally specified models," *J. Royal Statist. Soc., Series B*, vol. 61, no. 1, pp. 251–264, 1999.
- [30] A. Pázman, "Probability distribution of the multivariate nonlinear least squares estimates," *Kibernetika (Prague)*, vol. 20, pp. 209–230, 1984.
- [31] J. G. Propp, and D. B. Wilson, "Exact sampling with coupled Markov chains and applications to statistical mechanics," *Random Structures and Algorithms*, vol. 9, pp. 223–252, 1996.
- [32] C. P. Robert, and G. Casella, *Monte Carlo Statistical Methods*, Springer texts in Statistics, Springer Verlag, 2nd Edition, 2004.
- [33] C. P. Robert, *Méthodes de Monte Carlo par Chaînes de Markov*, Statistique Mathématique et Probabilité, Economica, France 1996.
- [34] L. Sandu, *Inversion des modèles paramétriques - Application aux mesures indirectes de températures*, Disserted Ph.D. thesis, Université Paris XI-SUPELEC, 1996.



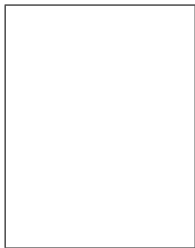
- [35] L. Sandu, J. Oksman, and G. Fleury, “Information criteria for the choice of parametric functions for measurement”, *IEEE Trans. Instrum. Meas.*, Vol 47, No. 4, pp. 920–924, 1998.
- [36] J. Shao, “Bootstrap model selection,” *J. of the American Statist. Assoc.*, Vol. 91, 434, pp. 655-665, 1996.
- [37] G. P. Terrell, “The maximal smoothing principle in density estimation,” *J. of the American Statist. Assoc.*, Vol. 85, pp. 470–477, 1990.
- [38] E. Walter, and L. Pronzato, *Identification de modèles paramétriques à partir de données expérimentales*, Masson, Paris, France 1994.
- [39] C. F. J. Wu, “Jackknife, Bootstrap and other resampling methods in regression analysis,” *Annals of Statist.*, Vol. 14, pp. 1261-1295, 1986.
- [40] A. M. Zoubir, “The Bootstrap and its application in signal processing”, *IEEE Signal Proc. Magazine*, Vol. 15, 1, pp. 56-76, 1998.



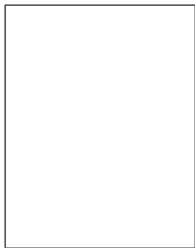
**José I. De la Rosa** was born in Zacatecas, Mexico, on March 19, 1972. He received the B.Sc. degree in electronics and telecommunications engineering from the Universidad Autónoma de Zacatecas, Mexico, in 1995, the M.Sc. degree in digital systems from the Instituto Politécnico Nacional (CITEDI), Mexico, in 1998, and the Ph.D. degree in signal processing and control from the Université de Paris-Sud, Orsay, France, in 2002. He was a Ph.D. student within the Measurement Department (now Department of Signal Processing and Electronic Systems) at SUPELEC. He is presently an Associate-Researcher at the Signal Processing Laboratory of the engineering faculty from the Universidad Autónoma de Zacatecas. His main interests are statistical signal processing, modelling and system identification, and signal and image processing in robotics.



**Gerardo Miramontes** was born in Zacatecas, Mexico, in 1956. He received the B.Sc. degree in electrical engineering from the Universidad Autónoma de Zacatecas, Mexico, in 1978, the M.Sc. degree in digital systems from the Instituto Politécnico Nacional (CITEDI), Mexico, in 2003. She was a M.Sc. student in stage within the Measurement Department of SUPELEC in 2001. Her main interests are statistical signal processing and Telecommunications.



**Gilles A. Fleury** was born in Bordeaux, France, on January 8, 1968. He received the B.Sc. degree from the École Supérieure d'Électricité (SUPELEC), Gif-sur-Yvette France, in 1990, the Ph.D. degree in signal processing from the Université de Paris-Sud, Orsay, France, in 1994, and his Habilitation a Diriger des Recherches (HDR) in 2003. He is presently a Professor within the Department of Signal Processing and Electronic Systems of SUPELEC. He has worked in the areas of inverse problems and optimal design. His current research interests include bioinformatics, optimal nonlinear modeling, and non uniform sampling.



**Marie-Eve Davoust** was born in Thorigny-sur-Marne, France, in 1948. She received the Ph.D. of Physics from the University of Paris 6, Paris, France, in 1978 and her Habilitation a Diriger des Recherches (HDR) from Universit Paris-Sud SUPELEC in 2005. She is currently a Professor at SUPELEC. She has been involved in various research topics in the signal processing. Her main interests are physical system modelling and signal processing for measurement purposes and microsystems computer aided design. She teaches courses on sensors and Non Destructive Evaluation.

## FIGURE AND TABLE CAPTIONS

Figure 1: Physical groove dimensions ( $l$  and  $d$ ) that may occur in ferromagnetic conductive pipes (inside).

Figure 2: Schematic illustration of the sensor pipe in the (RFEC) inspection problem.

Figure 3: The full line is the data fitted by the model using the mean parameter estimates from PMC-NLS<sup>2</sup>, (...) from Gibbs (MCMC<sup>2</sup>) and (- - -) from Gibbs M-H (MCMC<sup>4</sup>) vs. the simulated data (circles).

Figure 4: The full line is the data fitted by the model using the mean parameter estimates from PMC-NLS<sup>1</sup>, the other lines represent data fitted when using the MEE estimators (tree kernels) vs. the analyzed data (circles  $y$  against  $x$ ).

Figure 5: MSE according to the fitting sense  $y$  vs.  $f(\hat{\theta}, x)$ .

Figure 6: MSE according to the measurement estimation  $m$  (reference values) vs.  $\hat{m}$ .

Figure 7: Hypothesis Test one.

Figure 8: Hypothesis Test two.

Figure 9: Hypothesis Test three.

Table 1: Several statistics obtained for  $\theta$  and for  $m$ , when using PMC-NLS and MCLHS-NLS ( $e_i \sim \mathcal{N}(t, \hat{\sigma}_{ML}^e)$ ,  $\hat{\sigma}_{ML} = 0.0139$ ).

Table 2: Several statistics obtained with bootstrap methods and the least squares estimator (NLS).

Table 3: Several statistics obtained with the different bootstrap schemes (statistical characterization and biased estimators).

Table 4: Several statistics obtained for  $\theta$  and for  $m$ , according to the Gibbs resampling schema (MCMC<sup>1</sup>) and to the Gibbs CFTP schema (MCMC<sup>2</sup>).

Table 5: Several statistics obtained for  $\theta$  and for  $m$ , according to the hybrid sampling MCMC<sup>3</sup> and MCMC<sup>4</sup>.

Table 6: Several statistics obtained for  $\theta$  and for  $m$ , according to the bootstrap-MEE proposed schemes.

Table 7: First estimated values for  $d$  and  $l$ , according to the different used estimators, and the MSE as a quality measure of the fitting  $y$  and the measurement estimation  $m$ .

Table 8: Mean values for  $d$  and  $l$ , according to the different used estimators, and the MSE as a quality measure of the fitting  $y$  and the measurement estimation  $m$  (1).

Table 9: Mean values for  $d$  and  $l$ , according to the different used estimators, and the MSE as a quality measure of the fitting  $y$  and the measurement estimation  $m$  (2).

Table 10: Test results obtained in evaluating gaussian behavior of errors  $\wp(e)$ , according to the different estimators used.

## FIGURES

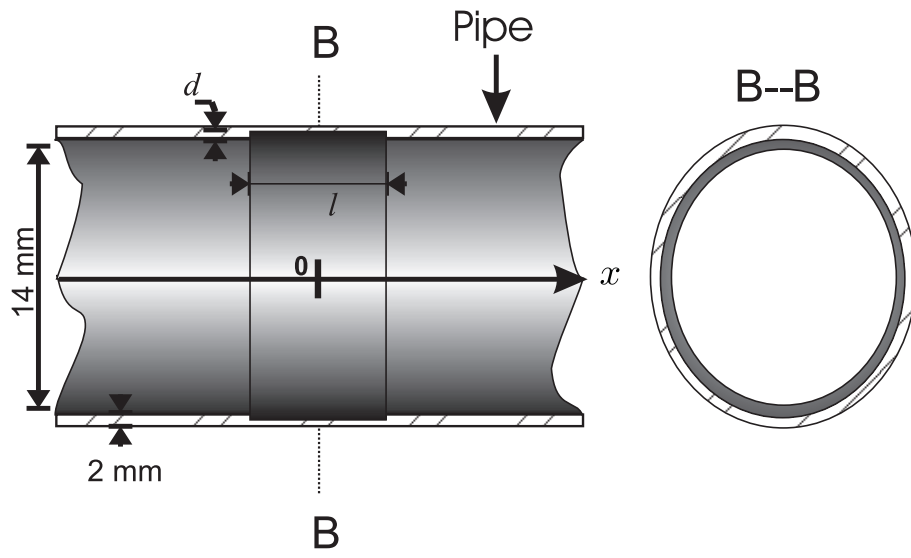


Fig. 1. Physical groove dimensions ( $l$  and  $d$ ) that may occur in ferromagnetic conductive pipes (inside).

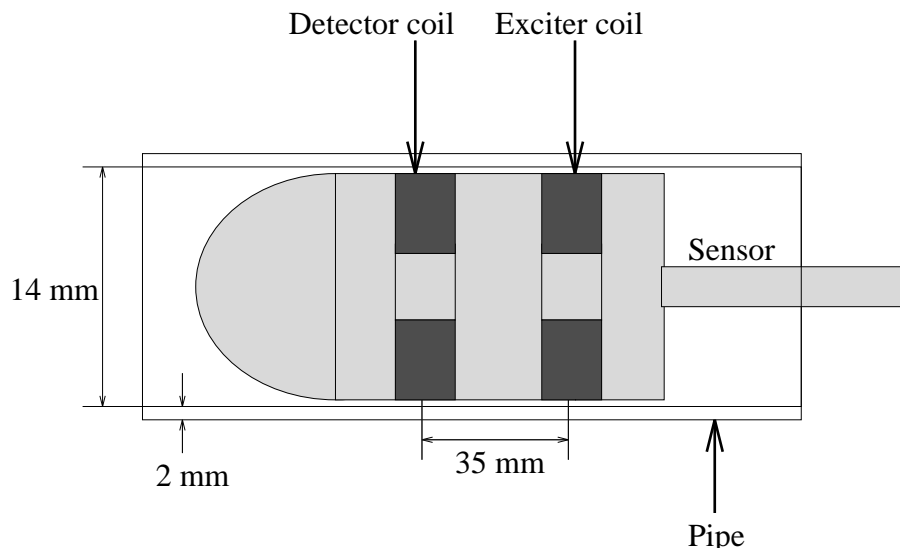


Fig. 2. Schematic illustration of the sensor pipe in the (RFEC) inspection problem.

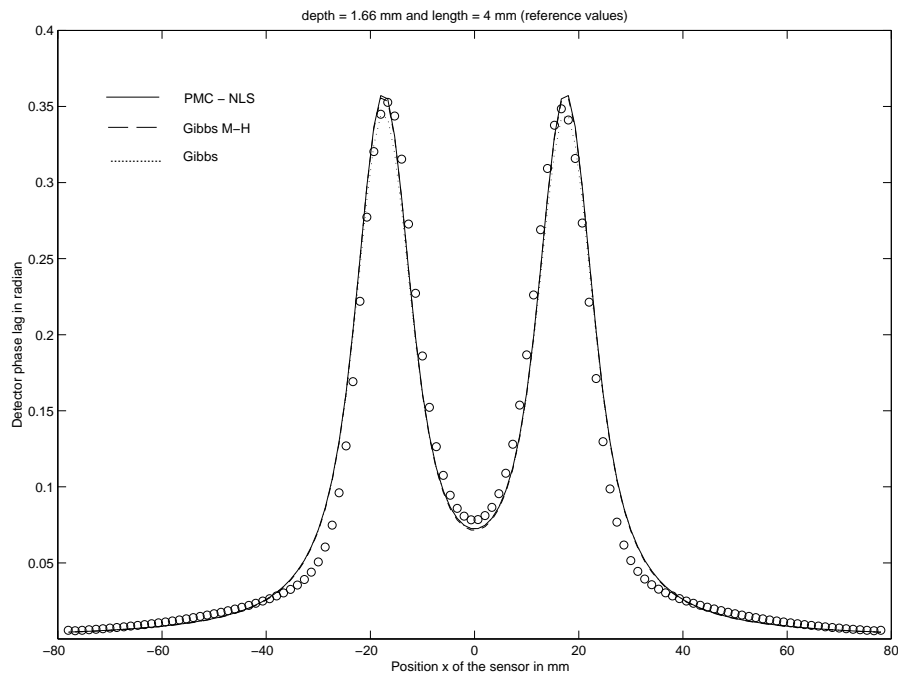


Fig. 3. The full line is the data fitted by the model using the mean parameter estimates from PMC-NLS<sup>2</sup>, (...) from Gibbs (MCMC<sup>2</sup>) and (- - -) from Gibbs M-H (MCMC<sup>4</sup>) vs. the simulated data (circles).

## TABLES

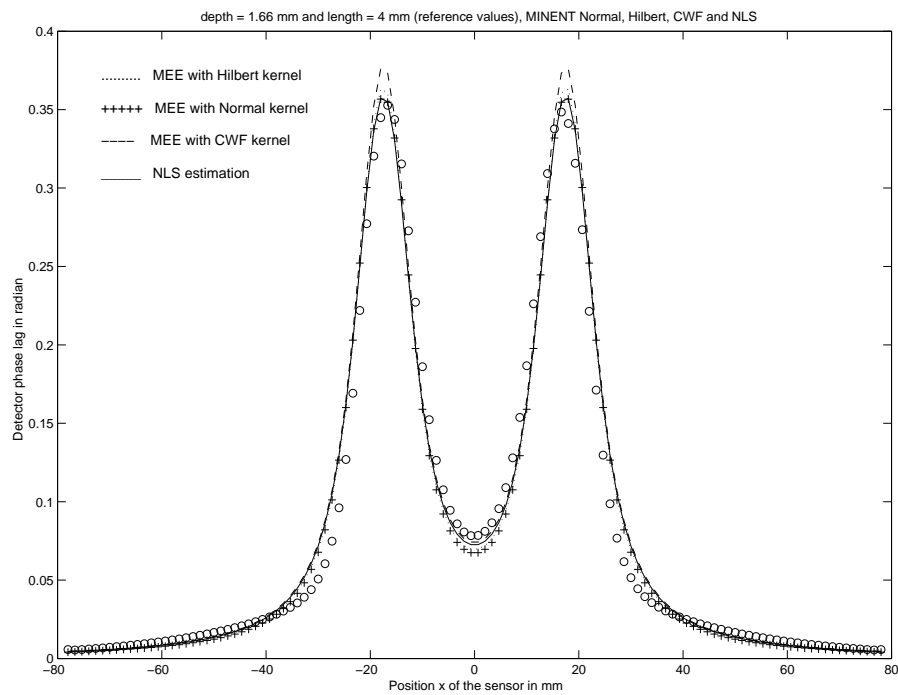


Fig. 4. The full line is the data fitted by the model using the mean parameter estimates from PMC-NLS<sup>1</sup>, the other lines represent data fitted when using the MEE estimators (tree kernels) vs. the analyzed data (circles  $y$  against  $x$ ).

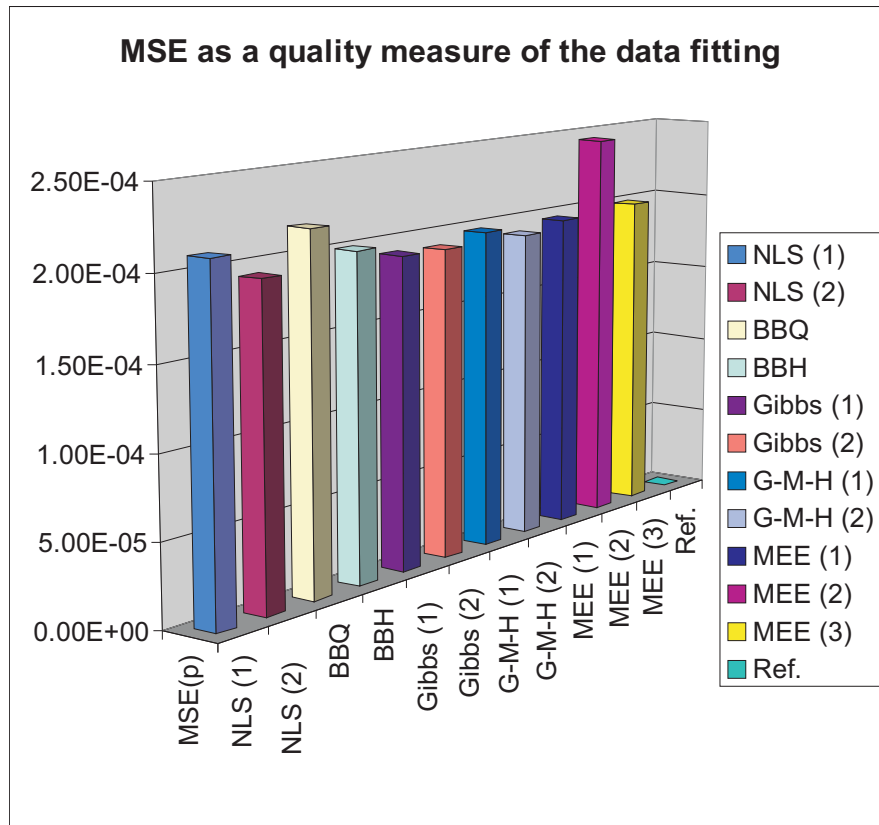


Fig. 5. MSE according to the fitting sense  $y$  vs.  $f(\hat{\theta}, x)$ .

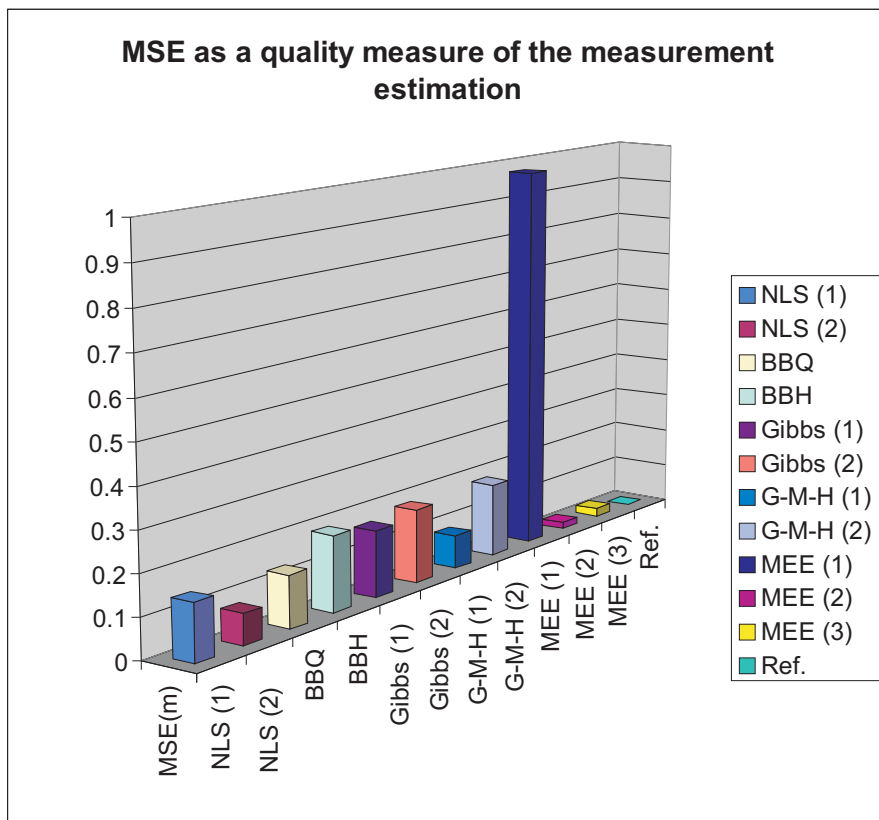


Fig. 6. MSE according to the measurement estimation  $m$  (reference values) vs.  $\hat{m}$ .

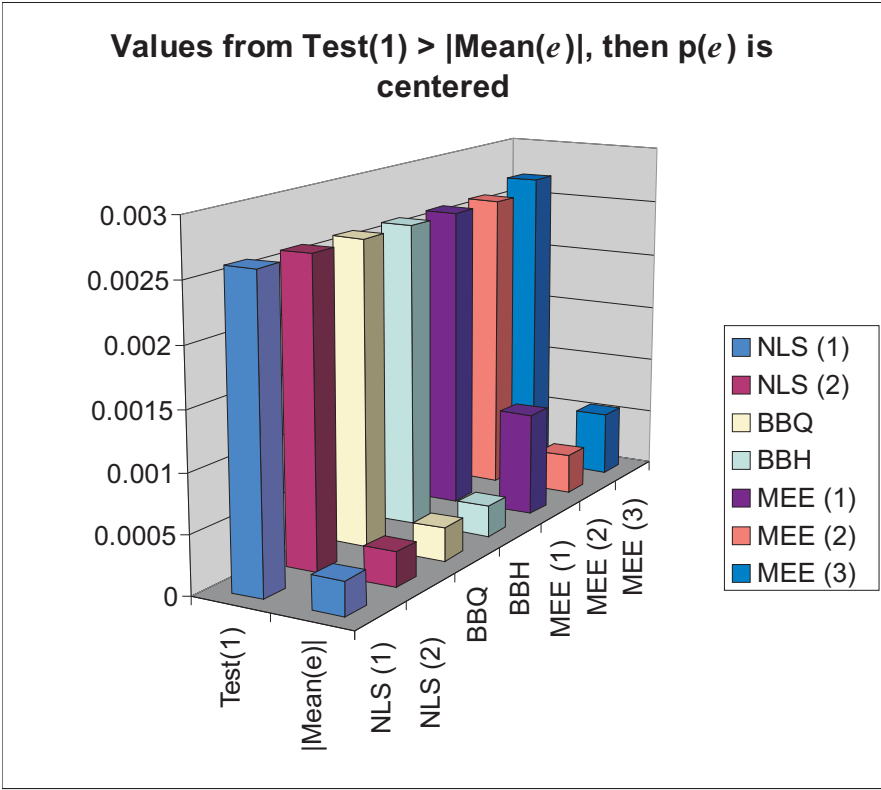


Fig. 7. Hypothesis Test one.

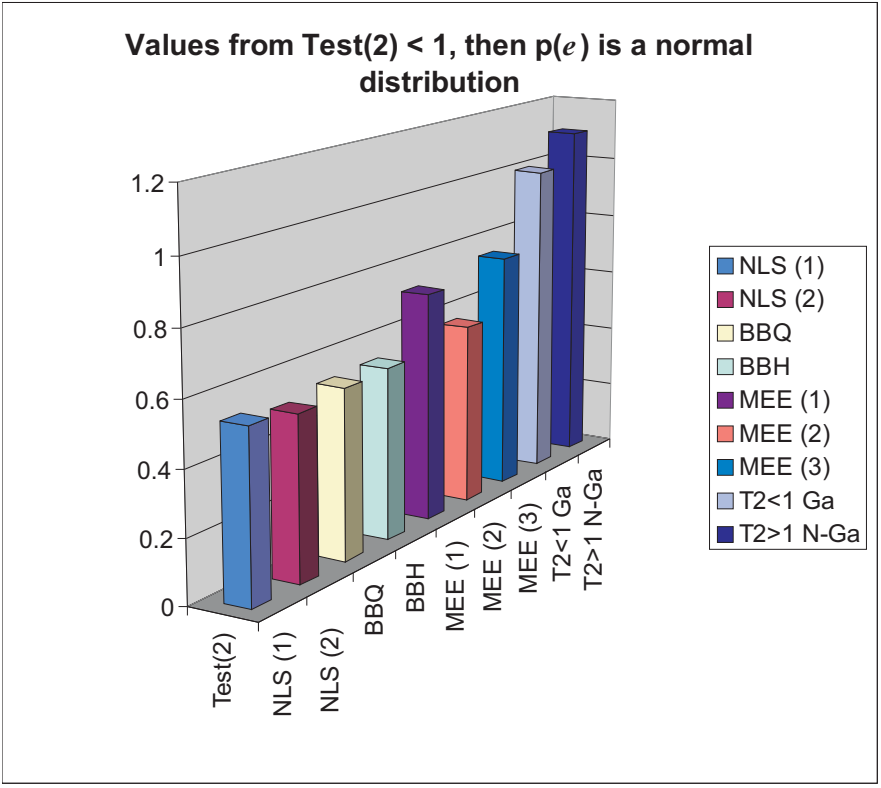


Fig. 8. Hypothesis Test two.

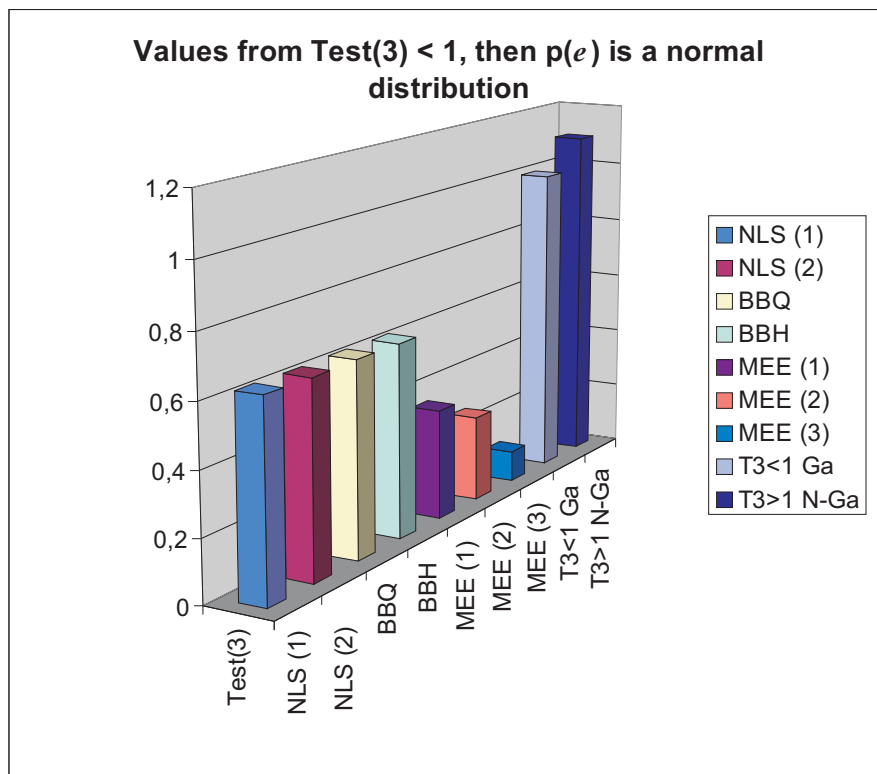


Fig. 9. Hypothesis Test three.



TABLE I

SEVERAL STATISTICS OBTAINED FOR  $\theta$  AND FOR  $m$ , WHEN USING PMC-NLS AND MCLHS-NLS ( $e_i \sim \mathcal{N}(0, \hat{\sigma}_{ML}^2)$ ,  $\hat{\sigma}_{ML} = 0.0139$ ).

MCP-NLS <sup>1</sup>	True	Mean	Median	Bias	Var	std
$\theta_1$	unknown	0.3081	0.2932	0.0141	0.0051	0.0716
$\theta_2$	unknown	0.1841	0.1839	4.0e-4	5.3193e-5	0.0073
$\theta_3$	unknown	3.5411	3.5991	0.0473	0.3485	0.5903
$d$	1.6600	1.6772	1.6803	0.0172	0.0208	0.1442
$l$	4.0000	4.5334	4.6159	0.5334	1.0122	1.0061
MCP-NLS <sup>2</sup>						
$\theta_1$	unknown	0.3097	0.3078	0.0148	0.0017	0.0409
$\theta_2$	unknown	0.1825	0.1815	1.2e-3	4.0761e-5	0.0064
$\theta_3$	unknown	3.4694	3.4584	0.1075	0.1687	0.4107
$d$	1.6600	1.7048	1.7292	0.0448	0.0132	0.1149
$l$	4.0000	4.3915	4.3574	0.3915	0.5153	0.7179
MCLHS-NLS <sup>1</sup>						
$\theta_1$	unknown	0.3104	0.2947	0.0164	0.0056	0.0749
$\theta_2$	unknown	0.1839	0.1837	1.4072e-4	5.2330e-5	0.0072
$\theta_3$	unknown	3.5221	3.5859	0.0663	0.3502	0.5918
$d$	1.6600	1.6818	1.6861	0.0218	0.0204	0.1430
$l$	4.0000	4.4998	4.5876	0.4998	1.0219	1.0109
MCLHS-NLS <sup>2</sup>						
$\theta_1$	unknown	0.3086	0.3073	0.0137	0.0018	0.0418
$\theta_2$	unknown	0.1829	0.1823	7.3078e-4	3.5790e-5	0.0060
$\theta_3$	unknown	3.4761	3.4493	0.1008	0.1826	0.4273
$d$	1.6600	1.6993	1.7242	0.0393	0.0125	0.1118
$l$	4.0000	4.4095	4.3522	0.4095	0.5516	0.7427

TABLE II

SEVERAL STATISTICS OBTAINED WITH BOOTSTRAP METHODS AND THE LEAST SQUARES ESTIMATOR (NLS).

BOOT-NLS <sup>1</sup>	True	Mean	Median	Bias	Var	std
$\theta_1$	unknown	0.3136	0.2959	0.0196	0.0064	0.0798
$\theta_2$	unknown	0.1833	0.1830	4.0e-4	6.6357e-5	0.0081
$\theta_3$	unknown	3.4981	3.5662	0.0903	0.3984	0.6312
$d$	1.6600	1.6894	1.6930	0.0294	0.0248	0.1575
$l$	4.0000	4.4600	4.5523	0.4600	1.1109	1.0540
BOOT-NLS <sup>2</sup>						
$\theta_1$	unknown	0.3095	0.2940	0.0155	0.0056	0.0746
$\theta_2$	unknown	0.1839	0.1838	2.0e-4	5.1572e-5	0.0072
$\theta_3$	unknown	3.5272	3.5899	0.0612	0.3472	0.5892
$d$	1.6600	1.6813	1.6826	0.0213	0.0202	0.1421
$l$	4.0000	4.5099	4.5977	0.5099	0.9986	0.9993

TABLE III

SEVERAL STATISTICS OBTAINED WITH THE DIFFERENT BOOTSTRAP SCHEMES (STATISTICAL CHARACTERIZATION AND BIASED ESTIMATORS).

BOOT-BBQ <sup>1</sup>	True	Mean	Median	Bias	Var	std
$\theta_1$	unknown	0.3163	0.2953	0.0223	0.0092	0.0958
$\theta_2$	unknown	0.1834	0.1832	0.0003	6.5982e-5	0.0081
$\theta_3$	unknown	3.4947	3.5722	0.0937	0.4052	0.6365
$d$	1.6600	1.6891	1.6915	0.0291	0.0248	0.1573
$l$	4.0000	4.4516	4.5730	0.4516	1.1574	1.0758
BOOT-BBQ <sup>2</sup>						
$\theta_1$	unknown	0.3111	0.2943	0.0171	0.0065	0.0807
$\theta_2$	unknown	0.1841	0.1839	0.0004	6.7850e-5	0.0082
$\theta_3$	unknown	3.5259	3.5857	0.0625	0.3988	0.6315
$d$	1.6600	1.6761	1.6826	0.0161	0.0257	0.1603
$l$	4.0000	4.5106	4.5901	0.5106	1.1414	1.0684
BOOT-BBH						
$\theta_1$	unknown	0.3009	0.2908	0.0068	0.0027	0.0516
$\theta_2$	unknown	0.1841	0.1837	4.0e-4	5.5635e-5	0.0075
$\theta_3$	unknown	3.5957	3.6480	0.0078	0.2598	0.5097
$d$	1.6600	1.6713	1.6824	0.0113	0.0193	0.1390
$l$	4.0000	4.6173	4.6838	0.6173	0.7975	0.8930

TABLE IV

SEVERAL STATISTICS OBTAINED FOR  $\theta$  AND FOR  $m$ , ACCORDING TO THE GIBBS RESAMPLING SCHEMA (MCMC<sup>1</sup>)  
AND TO THE GIBBS CFTP SCHEMA (MCMC<sup>2</sup>).

MCMC <sup>1</sup>	True	Mean	Median	Bias	Var	std
$\theta_1$	unknown	0.3001	0.2992	0.0052	0.0024	0.0491
$\theta_2$	unknown	0.1783	0.1784	0.0054	6.7908e-5	0.0082
$\theta_3$	unknown	3.5140	3.5245	0.0629	0.3303	0.5747
$d$	1.6600	1.7499	1.7534	0.0899	0.0197	0.1402
$l$	4.0000	4.5707	4.4967	0.5754	0.9853	0.9926
MCMC <sup>2</sup>						
$\theta_1$	unknown	0.2991	0.2982	0.0042	0.0028	0.0526
$\theta_2$	unknown	0.1787	0.1788	0.0050	7.7830e-5	0.0088
$\theta_3$	unknown	3.5256	3.5344	0.0513	0.3796	0.6161
$d$	1.6600	1.7436	1.7483	0.0174	0.0227	0.1509
$l$	4.0000	4.6035	4.5226	0.5568	1.1421	1.0704

TABLE V

SEVERAL STATISTICS OBTAINED FOR  $\theta$  AND FOR  $m$ , ACCORDING TO THE HYBRID SAMPLING MCMC<sup>3</sup> AND MCMC<sup>4</sup>.

MCMC <sup>3</sup>	True	Mean	Median	Bias	Var	std
$\theta_1$	unknown	0.3103	0.3010	0.0154	0.0024	0.0494
$\theta_2$	unknown	0.1834	0.1822	3.0e-4	5.0648e-5	0.0071
$\theta_3$	unknown	3.4752	3.5049	0.1017	0.2327	0.4824
$d$	1.6600	1.6917	1.7156	0.0317	0.0173	0.1316
$l$	4.0000	4.4100	4.4516	0.4100	0.7123	0.8440
MCMC <sup>4</sup>						
$\theta_1$	unknown	0.2979	0.2924	0.0030	0.0017	0.0407
$\theta_2$	unknown	0.1846	0.1844	9.0e-4	4.1251e-5	0.0064
$\theta_3$	unknown	3.5832	3.5925	0.0063	0.1849	0.4300
$d$	1.6600	1.6684	1.6770	0.0084	0.0139	0.1180
$l$	4.0000	4.6070	4.6226	0.6070	0.5738	0.7575

TABLE VI

SEVERAL STATISTICS OBTAINED FOR  $\theta$  AND FOR  $m$ , ACCORDING TO THE BOOTSTRAP-MEE PROPOSED SCHEMES.

BOOT-MEE <sup>1</sup>	True	Mean	Median	Bias	Var	std
$\theta_1$	unknown	0.2610	0.2576	0.0068	8.7766e-4	0.0296
$\theta_2$	unknown	0.1969	0.1967	0.0021	4.6103e-5	0.0068
$\theta_3$	unknown	4.0245	4.0384	0.1253	0.1355	0.3681
$d$	1.6600	1.4458	1.4534	0.2142	0.0210	0.1449
$l$	4.0000	5.3810	5.3993	1.3810	0.4320	0.6573
BOOT-MEE <sup>2</sup>						
$\theta_1$	unknown	0.3512	0.3295	0.0336	0.0079	0.0887
$\theta_2$	unknown	0.1832	0.1829	0.0028	4.2665e-5	0.0065
$\theta_3$	unknown	3.1579	3.2417	0.2334	0.3054	0.5526
$d$	1.6600	1.7244	1.7269	0.0644	0.0180	0.1343
$l$	4.0000	3.8302	3.9569	0.1698	0.8466	0.9201
BOOT-MEE <sup>3</sup>						
$\theta_1$	unknown	0.3138	0.3166	0.0048	2.8911e-4	0.0170
$\theta_2$	unknown	0.1883	0.1879	0.0013	2.4532e-5	0.0050
$\theta_3$	unknown	3.3842	3.3476	0.0436	0.0366	0.1913
$d$	1.6600	1.6378	1.6456	0.0222	0.0055	0.0739
$l$	4.0000	4.2104	4.1447	0.2104	0.1102	0.3320

TABLE VII

FIRST ESTIMATED VALUES FOR  $d$  AND  $l$ , ACCORDING TO THE DIFFERENT USED ESTIMATORS, AND THE MSE AS A QUALITY MEASURE OF THE FITTING  $y$  AND THE MEASUREMENT ESTIMATION  $m$ .

$\backslash$	NLS <sup>1</sup>	NLS <sup>2</sup>	BBQ	BBH	MEE <sup>1</sup>	MEE <sup>2</sup>	MEE <sup>3</sup>
MSE $_{\theta}$	1.8805e-4	1.8805e-4	1.8805e-4	1.8805e-4	1.9313e-4	1.9499e-4	2.0615e-4
$\hat{d}$	1.6835	1.6851	1.6835	1.6837	1.4966	1.6750	1.6269
$\hat{l}$	4.5992	4.5796	4.5992	4.5984	5.1456	4.1440	4.1143
MSE $_m$	0.1798	0.1683	0.1798	0.1793	0.6695	0.0105	0.0071

TABLE VIII

MEAN VALUES FOR  $d$  AND  $l$ , ACCORDING TO THE DIFFERENT USED ESTIMATORS, AND THE MSE AS A QUALITY MEASURE OF THE FITTING  $\mathbf{y}$  AND THE MEASUREMENT ESTIMATION  $\mathbf{m}$  (1).

$\backslash$	NLS <sup>1</sup>	BBQ	BBH	MEE <sup>1</sup>	MEE <sup>2</sup>	MEE <sup>3</sup>
MSE $_{\theta}$	2.0939e-4	2.1695e-4	1.9955e-4	1.9653e-4	2.4432e-4	1.9913e-4
$\bar{d}$	1.6772	1.6761	1.6713	1.4458	1.7244	1.6378
$\bar{l}$	4.5334	4.5106	4.6173	5.3810	3.8302	4.2104
MSE $_{\mathbf{m}}$	0.1424	0.1305	0.1906	0.9765	0.0165	0.0224

TABLE IX

MEAN VALUES FOR  $d$  AND  $l$ , ACCORDING TO THE DIFFERENT USED ESTIMATORS, AND THE MSE AS A QUALITY MEASURE OF THE FITTING  $\mathbf{y}$  AND THE MEASUREMENT ESTIMATION  $\mathbf{m}$  (2).

$\backslash$	NLS <sup>2</sup>	MCMC <sup>1</sup>	MCMC <sup>2</sup>	G-M-H <sup>1</sup>	G-M-H <sup>2</sup>	Ref.
MSE $_{\theta}$	1.9332e-4	1.9192e-4	1.9147e-4	1.9715e-4	1.9124e-4	-
$\bar{d}$	1.7048	1.7499	1.7436	1.6917	1.6684	1.6600
$\bar{l}$	4.3915	4.5707	4.6035	4.4100	4.6070	4.0000
MSE $_{\mathbf{m}}$	0.0776	0.1669	0.1856	0.0846	0.1843	-

TABLE X

TEST RESULTS OBTAINED IN EVALUATING GAUSSIAN BEHAVIOR OF ERRORS  $\varphi(\mathbf{e})$ , ACCORDING TO THE DIFFERENT ESTIMATORS USED.

$\backslash$	NLS <sup>1</sup>	NLS <sup>2</sup>	BBQ	BBH	MEE <sup>1</sup>	MEE <sup>2</sup>	MEE <sup>3</sup>
$ \hat{m}_e $	2.8297e-4	2.9527e-4	2.8297e-4	2.6684e-4	8.8580e-4	3.4272e-4	5.4981e-4
Test <sub>1</sub>	0.0026	0.0026	0.0026	0.0026	0.0026	0.0026	0.0027
Test <sub>2</sub>	0.5315	0.5087	0.5315	0.5395	0.7202	0.5694	0.7497
Test <sub>3</sub>	0.6243	0.6207	0.6243	0.6221	0.3547	0.2739	0.0999